

University of Arkansas, Fayetteville

ScholarWorks@UARK

Biomedical Engineering Undergraduate Honors
Theses

Biomedical Engineering

5-2020

Theoretical Applicability of CRISPR-Mediated Exon Skipping in Duchenne Muscular Dystrophy

Andrew Greek

Follow this and additional works at: <https://scholarworks.uark.edu/bmeguht>



Part of the [Other Biomedical Engineering and Bioengineering Commons](#)

Citation

Greek, A. (2020). Theoretical Applicability of CRISPR-Mediated Exon Skipping in Duchenne Muscular Dystrophy. *Biomedical Engineering Undergraduate Honors Theses* Retrieved from <https://scholarworks.uark.edu/bmeguht/87>

This Thesis is brought to you for free and open access by the Biomedical Engineering at ScholarWorks@UARK. It has been accepted for inclusion in Biomedical Engineering Undergraduate Honors Theses by an authorized administrator of ScholarWorks@UARK. For more information, please contact ccmiddle@uark.edu.

Theoretical Applicability of CRISPR-Mediated Exon Skipping in Duchenne Muscular Dystrophy

Honors Thesis by Andrew Greek

Department of Biomedical Engineering

College of Engineering

University of Arkansas

Faculty Mentor: Dr. Christopher Nelson

Honors Coordinator: Dr. Kyle Quinn

Table of Contents

1. Abstract	3
2. Introduction	4
2.1. Muscular Dystrophy	4
2.2. CRISPR and Exon Skipping	5
3. Materials and Methods	5
3.1. Data Formatting	5
3.2. Mapping the Mutations	6
3.3. Modelling Exon Skipping	8
4. Results	11
5. Discussion	15
6. Future Directions	16
7. Conclusion	17
8. Acknowledgements	17
References	18

1. Abstract

Duchenne muscular dystrophy (DMD) affects approximately 1 in 5,000 males. DMD results from genetic mutations within the gene that interrupts the open reading frame and prematurely truncates the protein. From a non-functional dystrophin protein, DMD results in serious muscle weakening and muscle wasting ultimately leading to death around the age of 26. In some cases of muscular dystrophy, a mutation can occur in areas of the gene that are less essential. In Becker muscular dystrophy (BMD), these mutations cause less significant consequences in phenotype as the dystrophin protein is still semi-functional. With gene editing techniques, it may be possible to reduce the consequences of the DMD phenotype. Exon skipping is a process that advertently skips specific and surrounding exons that have been harmfully mutated. This process could potentially be used to skip specific exons in the DMD gene necessary to bring the RNA transcript back into the correct reading frame. This study focuses on the theoretical application of exon skipping in DMD. Using data of all the recorded mutations up to now, the study looks at the separate effects skipping each exon has on the three different types of mutations: deletions, duplications, and small rearrangements. Exon skipping was modeled computationally. Under an important operational assumption, the results were encouraging. Theoretically, skipping the top eight most effective exons would correctly reframe just above 50% of the total mutations. Skipping exon 51 alone would restore the reading frame in 14% of patients. This study focused primarily on the results of skipping only one exon at a time. While there is still much more work to do to find the true applicability of exon skipping in DMD, the results provide a solid framework and guide to what comes next.

2. Introduction

2.1 Muscular Dystrophy

Duchenne muscular dystrophy (DMD) is an X-linked recessive genetic disorder found in approximately 1 in 5,000 males (NIH, 2016). The genetic disorder is caused by mutations in the DMD gene leading to premature truncation of the gene. The DMD gene is 79 exons long and is responsible for the production of the dystrophin protein. When complete, this protein is used to strengthen muscle fibers and protect them from injury as muscles contract and relax (NIH, 2017). When a mutation causes the gene to be non-functional, the consequences in the corresponding phenotype can be devastating. Motor skills suffer initially as muscle weakness and muscle wasting occurs in the voluntary muscles. As it progresses, this degenerative disease eventually leads to the loss of ambulation of those affected. Involuntary and cardiac muscles are also weakened leading to the eventual death of the patient. The average life expectancy associated with DMD is 26 years.

In some cases of muscular dystrophy, a mutation can occur in areas of the gene that are less essential. In Becker muscular dystrophy (BMD), these mutations cause less significant consequences in phenotype as the dystrophin protein is still semi-functional. With a significantly longer life expectancy and greater quality of living, BMD clearly results in better conditions for patients than does DMD. This provided the basis for a potential DMD treatment. Becker muscular dystrophy shows that the DMD gene can still produce functional dystrophin proteins when the mutation does not result in catastrophic frameshifts, premature truncation, or occur in highly important areas of the gene. With this information, a gene editing approach has been proposed to reframe the exons of a DMD gene to prevent a typical DMD phenotype.

2.2. CRISPR and Exon Skipping

A common approach to gene editing has been with the use of clustered regularly interspaced short palindromic repeats or CRISPR. CRISPR is a bacterial immune system capable of producing double strand breaks in target DNA. A clinical approach for CRISPR has been producing double-strand breaks in a target gene and allowing the cell to repair to form small insertions or deletions (indels). These indels can be used to knock out target exons in a manner similar to antisense-mediated exon skipping.

In Duchenne muscular dystrophy, there is often an exon or group of exons that are mutated. This disrupts the open-reading frame and often results in a non-functional protein (Aartsma-Rus, 2006). Examples from Becker muscular dystrophy, however, show that the protein can still be functional even while missing some of its non-essential regions. With this concept, exon skipping provides a hopeful tool for DMD by skipping necessary exons to bring the RNA transcript back into the correct reading frame. This study aims to show the theoretical applicability of using this technique in regards to how many currently recorded cases of DMD would undergo a reading-frame restoration through the skipping of a single exon.

3. Materials and Methods

3.1. Data Formatting

All of the data was obtained from the UMD website (UMD, n.d.). 7,150 records are broken up into 1,733 different mutations. Information pertaining to the protein nomenclature, cDNA nomenclature, the exon and codon affected, structure within the gene, type of rearrangement, and number of records for each mutation is all provided. The data was copied

into Microsoft Excel for further manipulation. Using basic Excel functions, the data was separated by mutation and sorted from most common to least common.

Protein nomenclature	cDNA Nomenclature	Exon	Codon	Structure	HCD	Rearrangement	Mutation type	Mutational event	# records
p.Glu2147AlafsX17	c.6439_6614del	45	2147	CRD-repeat #17		Large rearrangement Deletion from exon 45 to 45	Fr.	Stop at 2163	317
p.Glu2147LeufsX9	c.6439_7309del	45	2147	CRD-repeat #17		Large rearrangement Deletion from exon 45 to 50	Fr.	Stop at 2155	268
p.Val2305LeufsX9	c.6913_7309del	48	2305	CRD-repeat #18		Large rearrangement Deletion from exon 48 to 50	Fr.	Stop at 2313	250
p.Glu2147LeufsX22	c.6439_7660del	45	2147	CRD-repeat #17		Large rearrangement Deletion from exon 45 to 52	Fr.	Stop at 2168	224
p.Glu2367LeufsX9	c.7099_7309del	49	2367	CRD-repeat #19		Large rearrangement Deletion from exon 49 to 50	Fr.	Stop at 2375	192

Figure 1: Each mutation was organized using the 9 categories shown above

3.2. Mapping the Mutations

The next step was to model the mutations within the 79 exons of the DMD gene. The exon structure of the gene was plotted out to the right of the mutational data. Exons 1-79 were used as column headers. There were three main types of mutations: deletions, duplications, and small rearrangements. For each mutation, the cell that was representing the exon or exons affected was shaded in. This was done using a binary number system. Each cell contained either a 0 or a 1. If it contained a 0, it was not shaded in. If it contained a 1, the cell was shaded representing a mutation of that exon.

[illegible]

Figure 2: A portion of the mutation map between exons 45 and 54.

In order to map the placement of each mutation, the data on the range of affected exons needed to be isolated. Most rows of the “Rearrangement” column included the range of exons which the mutation affected (“from exon # to #). Each row of the “Rearrangement” column was searched for the word “exon” (Eq 1). The following numeric value was extracted from the cell (Eq 2). Again the column was searched up to the number of the ending exon and then extracted out as well using similar search, mid, and value functions (Eq 3 and 4). Once the exon numbers were isolated, each mutation was sifted through exon by exon and assigned a 0 or a 1 for each exon on the exon map. This was done using the product of two “If” logic functions (Eq 5). If the exon number was greater than or equal to the first exon number isolated in that row, then it was assigned a 1. If the exon number was less than or equal to the last exon number isolated in that

row, then it was also assigned a 1. All other exons for each mutation was given a value of 0. Simple conditional formatting shaded in the cells with a value of 1. This allowed the affected area of the gene to be clearly represented with shaded cells for each mutation.

Formula 1: =SEARCH("exon ",G4,1)

Formula 2: =VALUE(MID(G4,L4+5,2))

Formula 3: =SEARCH(" to ",G4,"1")

Formula 4: =VALUE(MID(G4,N4+4,2))

Formula 5: =IF(\$\$3>=\$M4,1,0)*IF(\$\$3<=\$O4,1,0)

3.3. Modelling Exon Skipping

After mapping the mutations on the excel sheet, it was necessary to look at the results of skipping individual exons. To do this, the structure and reading frame of each exon needed to be represented. Common literature often depicts this by using exons with shaped ends corresponding to the next exon. To model this compatibility, each exon was assigned two reading-frame numbers, 1-3, with each number corresponding to the shape of both the 5' and 3' end of the exon.



Figure 3: Normal reading frame of the DMD gene from exon 68 to exon 75 (Muscular Dystrophy UK, 2015)



Figure 4: Deleted exon 52 causing a frameshift and prematurely truncating the protein
(Muscular Dystrophy UK, 2015)

	1 2	2 1	1 1	1 1	1 1	1 3	3 2
Exons -->	1	2	3	4	5	6	7

Figure 5: Illustration of reading frame numbers assigned to the individual exons

Modelling the exon skipping approach was different depending on the type of mutation. For the deletions, the approach needed to be broken down in two parts: skipping the exon before the mutation and after the mutation. Combining these two would result in the theoretical applicability of skipping each exon. For the duplications, this study focused on single exon duplications as multiple exon deletions would likely require skipping more than one exon (Niks, 2017). Finally, for the small rearrangements, the exon skipped was the one with the rearrangement and thus causing an out-of-frame mutation. An important operating assumption made initially was that this study was going to neglect where in the gene the mutation and skipping was taking place. Thus, it looked at only restoring the reading frame and saved analyzing the phenotypic consequences for a future study.

Beginning with skipping one exon before a deletion, each row was searched to find the first affected exon and then extracted the exon two before that (Eq 6). The 3' reading-frame number assigned to that exon was isolated and saved for later comparison (Eq 7 and 8). Next, the same row was searched for the last affected exon and the exon after was extracted (Eq 9). The 5'

reading-frame number assigned to that exon was isolated and compared to the 3' reading-frame number of the first isolated exon (Eq 7 and 8). The two reading-frame numbers were analyzed using a simple “If” logical formula: if the two numbers were equal, then that row would be labelled “true”. If the two numbers were not equal, then that row would be labelled “false” (Eq 10). This process was repeated again but instead of skipping the exon before the deletion, the exon after the deletion was skipped (Eq 7, 8, 10, 11, and 12). Of the results that were labelled “true”, the exon skipped and the number of records of that mutation were organized into a pivot table and then graphed accordingly.

Formula 6: =MATCH(1,U4:CU4,0) -2

Formula 7: =INDEX(\$S\$2:\$CS\$3,1,CU4)

Formula 8: =MID(CV4,4,4)

Formula 9: =LOOKUP(2,1/(S4:CS4<>0),\$S\$3:\$CS\$3)+1

Formula 10: =IF(CW4=CZ4,"true", "false")

Formula 11: =MATCH(1,U4:CU4,0) -1

Formula 12: = LOOKUP(2,1/(S4:CS4<>0),\$S\$3:\$CS\$3)+2

For the duplications, a similar strategy was used. Each row was searched for the two exons on either side of the duplicated exon (Eq 9 and 11). The number of exons duplicated were calculated by subtracting the two exons on either side. Since the study is looking at duplications only affecting one exon, mutations that contained duplications longer than a single exon were filtered out. A pivot table was created consisting of the number of records from each skipped exon.

The process used for the small arrangements was only slightly different than what was used for the deletions. The two exons on the 5' and 3' sides of the affected area were isolated using equations 9 and 11. The two reading-frame numbers assigned to the isolated exons were extracted and compared in similar fashion to the deletion process (Eq 7, 8, and 10). Once each mutation was labelled “true” or “false” based on the restoration of the reading frame, all of the true results were organized into a pivot table displaying the number of records for each skipped exon.

4. Results

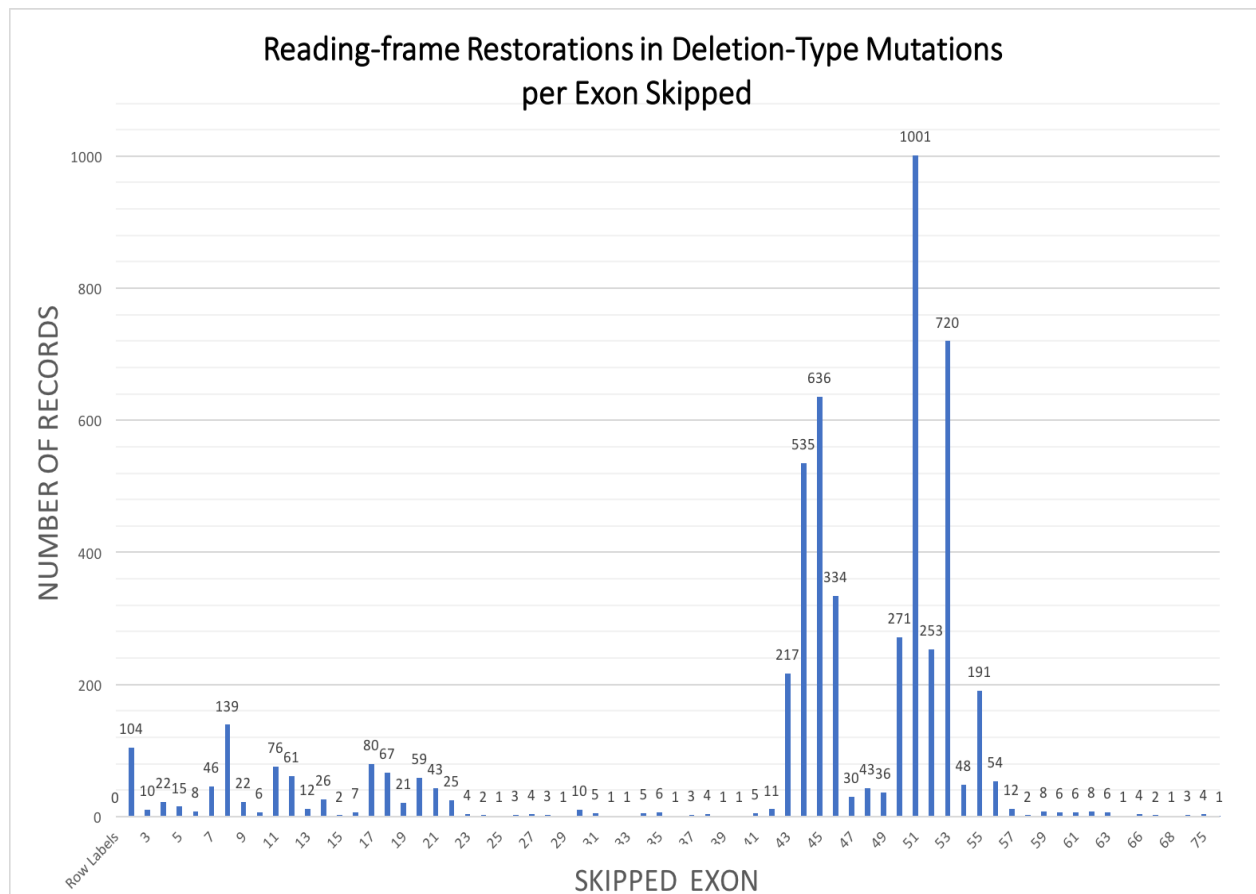


Figure 6: This graph depicts the amount of recorded deletion-type mutation cases of DMD that would theoretically have their reading-frame restored through a single exon skip. Skipping each individual exon, excluding the first and the last (1 and 79), yielded different results depending on where the deletions were occurring within the gene.

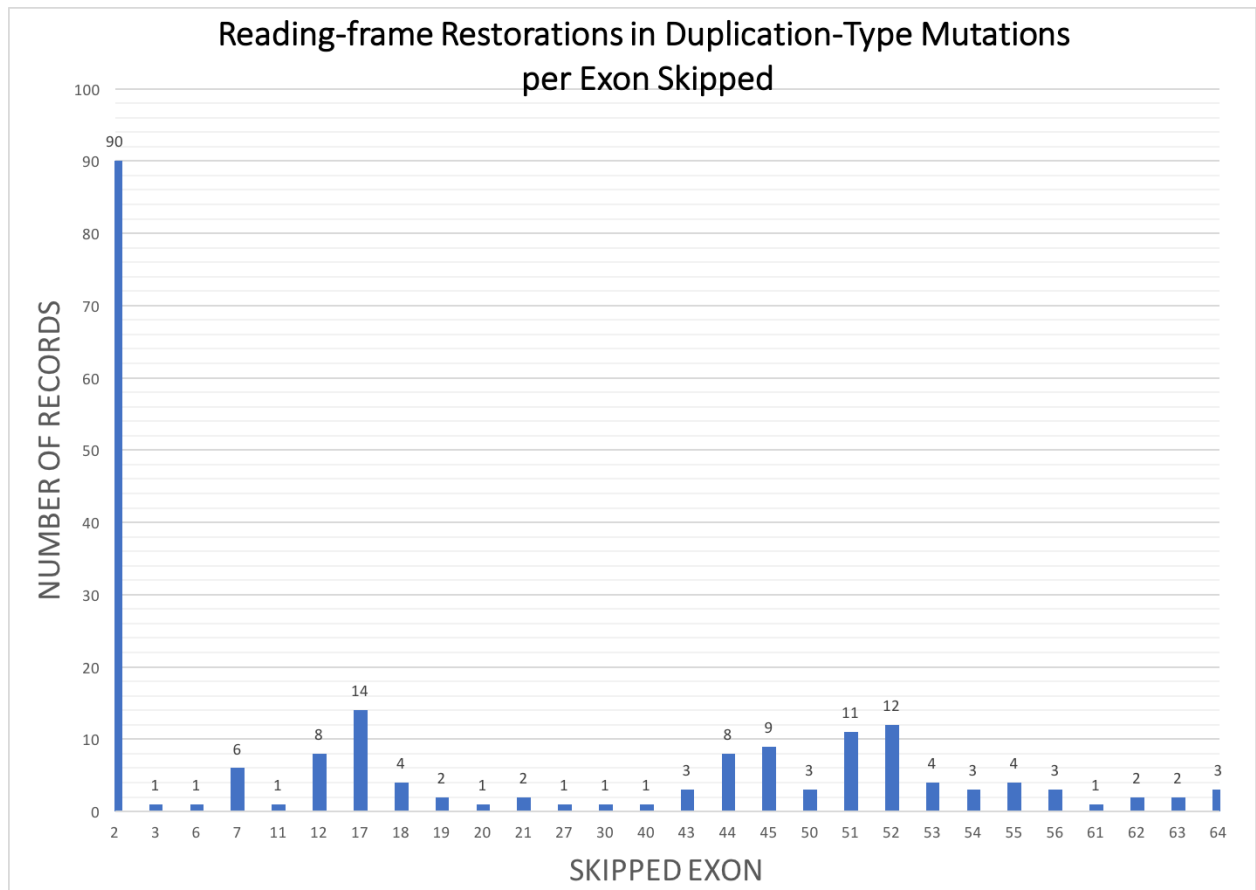


Figure 7: This graph shows the amount of recorded duplication-type mutation cases of DMD that would theoretically have their reading-frame restored through a single exon skip. Skipping each individual exon, excluding the first and the last (1 and 79), yielded different results depending on where the deletions were occurring within the gene. With duplications, only one of the two copies of an exon was skipped.

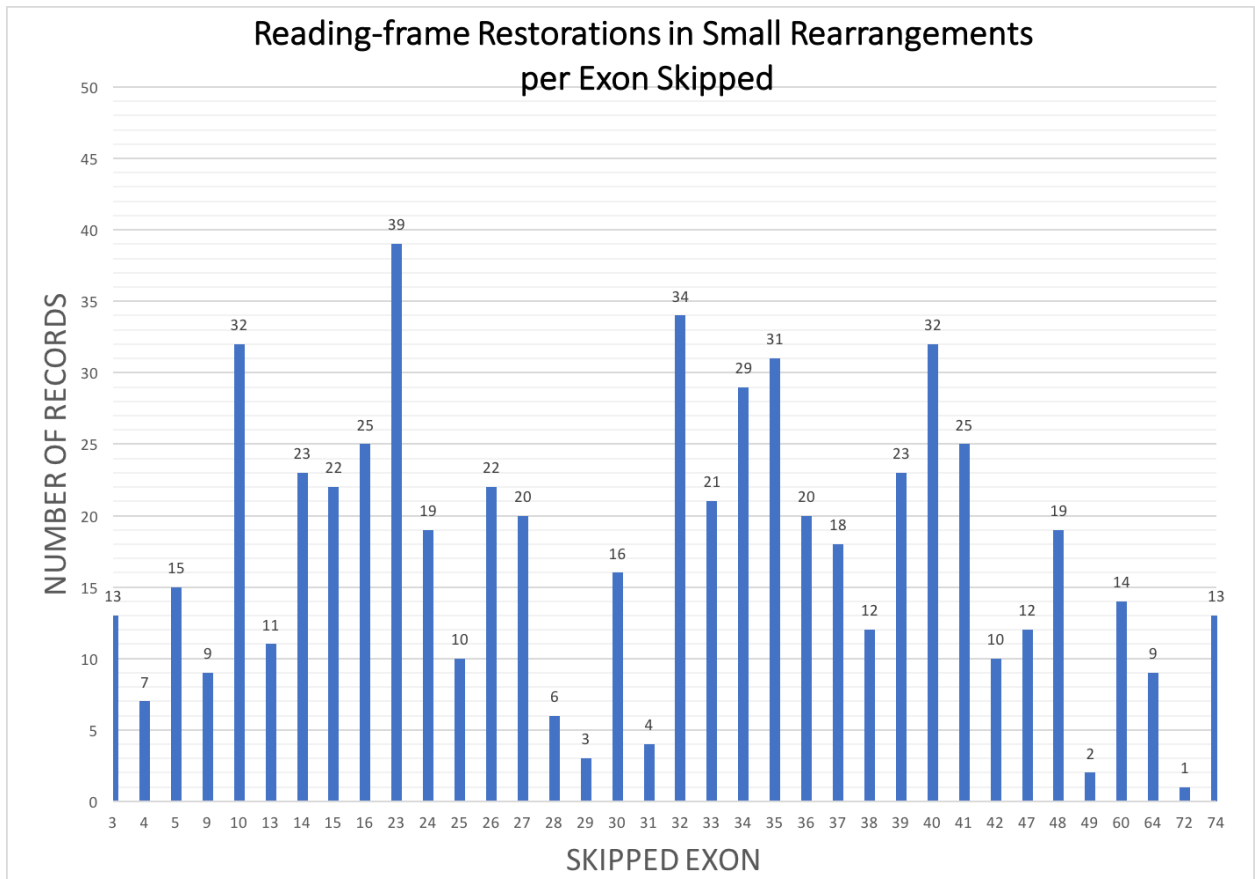


Figure 8: The above graph illustrates the amount of small rearrangement mutation cases that would theoretically have their reading-frame restored through a single exon skip.

Skipping each individual exon, excluding the first and the last (1 and 79), yielded different results depending on where the deletions were occurring within the gene. This study focused on the mutations with small rearrangements in one exon.



Figure 9: This figure represents every recorded case of DMD, regardless of the mutation type, that would theoretically have their reading-frame restored through a single exon skip. Skipping each individual exon, excluding the first and the last (1 and 79), yielded different results depending on where the deletions were occurring within the gene. This graph combined the results of figures 6, 7, and 8.

5. Discussion

The four graphs above show the theoretical results of exon skipping in three different types of mutations. Figure 6 shows the results of exon skipping for deletions of at least one exon. This group contained the largest number of records, 4903, and consisted of 506 different deletions. The graph also illustrates a hotspot for deletions between exons 43 and 55, thus skipping exons within this range affects the largest number of cases. Seen in Figure 6, exon 51 theoretically restores the reading-frame in 1001 cases out of 4903 deletions and 7150 total mutations for about 14% of all DMD mutations.

Figure 7 shows the results from the application of exon skipping to duplication mutations in DMD. The number of one exon duplications was significantly less than deletions; however, there is still a slight cluster of cases between exons 2 and 7 and also exons 43 and 55. On the other end of the gene, exon 2 is a commonly duplicated exon. Luckily, a deletion of exon 2 restores the genetic reading-frame in 90 cases. While there are six times more deletions than duplications, the strategy for exon skipping in duplication mutations is less straightforward. Single exon duplications require only one copy of the mutated exon to be skipped: either the original or duplicated exon. At the same time if both copies of the exon are skipped, it can lead to an out-of-frame transcript. Mutations involving multiple exon duplications often require multiple exons to be skipped and can be very challenging. Exploring other gene editing strategies would potentially yield better results for multiple exon duplications.

Figure 8 graphs the results from individual exon skipping in the small rearrangement mutations found in DMD. The results from this group of mutations were relatively even across the 79 exons. Two slight peaks can be seen from exons 14 to 23 and from exons 32 to 41. In single exon rearrangements, the reading-frame can often be restored by skipping the exon that

has been rearranged and thus has caused an out-of-frame transcript. Skipping exon 23 would produce the highest single exon benefit for small arrangement mutations affecting 39 cases.

Figure 9 shows the final count of restored cases for each exon in all DMD mutations. It closely resembles the shape of the graph in figure 6 as most of the total mutations are deletions. Exon 51 still reframes the largest number of cases at 1012, while exons 45 and 53 also contribute to restoring 645 and 724, respectively. There are two large clusters of highly influential exons from exon 43 to 46 and then from 50 to 53. These eight exons combine to theoretically reframe 3,797 cases, approximately 53.1% of the total 7150 cases.

6. Future Directions

An extension of this study is necessary in order to understand the complete applicability of exon skipping in DMD. This study looked only at single exon skips and does not address the possibility of skipping multiple consecutive exons at once. While the ability to skip multiple exons would presumably be of much greater difficulty, it would theoretically allow for a larger portion of mutations to be reframed. A study over this topic would likely require more intense programming.

This study was under the operational assumption that the functionality of the protein would not be affected regardless of where the exon deletions took place. Instead, it looked strictly at restoring the reading-frame. In reality, there are multiple regions of the gene that encode fundamental parts of the dystrophin protein and thus cannot not afford to be skipped; however, the hotspot region is usually considered fairly dispensable. A future study would look at the effects of this concept on this technique's applicability.

7. Conclusion

Overall, the data suggests gene editing to be very applicable to Duchenne muscular dystrophy. With only skipping one exon at a time, a total of just over 50% of all the recorded mutations could have a shift back into the correct reading-frame. In particular, skipping exon 51 alone was calculated to reframe 14% of the total mutations. Exon skipping holds the most promise for deletions largely because deletions make up the majority of the mutations; however, exon skipping for duplications and small rearrangements create other difficulties as well. This study has demonstrated the theoretical applicability of CRISPR-mediated exon skipping in Duchenne muscular dystrophy based on restorations in the reading frame. While providing groundwork for an extension of this study, it has highlighted a few key exons which, if targeted, have the potential to help many people suffering from DMD. While these numbers are theoretical and are expected to decrease slightly as other factors (exon location and technology limitations) are considered, this research has effectively illustrated its potential.

8. Acknowledgements

I would like to extend a special thank you to Dr. Christopher Nelson. This research project would not have been possible without his guidance and expertise. I would also like to thank the Biomedical Engineering department for the knowledge and direction it has provided over the last four years. Lastly, I am grateful to the Honors College at the University of Arkansas for allowing me to conduct this research in the form of an Honors Thesis.

References

Aartsma-Rus, A., & Van Deutekom, J. C. T. (2006). Entries in the Leiden Duchenne muscular dystrophy mutation database: An overview of mutation types and paradoxical cases that confirm the reading-frame rule. *Muscle & Nerve*, 34(2), 135–144.

Muscular Dystrophy UK. (2015). What is exon skipping and how does it work? Retrieved April 20, 2020, from <https://www.musculardystrophyuk.org/progress-in-research/background-information/what-is-exon-skipping-and-how-does-it-work/>

NIH, National Institute of Neurological Disorders and Stroke. (2016, March 4). Muscular Dystrophy: Hope Through Research. Retrieved from https://web.archive.org/web/20160930165657/http://www.ninds.nih.gov/disorders/md/detail_md.htm#180483171

Niks, E. H., & Aartsma-Rus, A. (2017). Exon skipping: a first in class strategy for Duchenne muscular dystrophy. *Expert Opinion on Biological Therapy*, 17(2), 225–236.

The UMD TREAT-NMD DMD mutations database. (n.d.). Retrieved April 20, 2020, from http://umd.be/TREAT_DMD/

NIH, U.S. National Library of Medicine. (2017, February). DMD gene - Genetics Home Reference - NIH. Retrieved April 20, 2020, from <https://ghr.nlm.nih.gov/gene/DMD>